One.III Reduced Echelon Form

Linear Algebra
Jim Hefferon

http://joshua.smcvt.edu/linearalgebra

Gauss-Jordan reduction

Pivoting

Example Here is an extension of Gauss's Method with some advantages.

Instead of doing back substitution, we continue using the row operations.

Pivoting

Example Here is an extension of Gauss's Method with some advantages.

Instead of doing back substitution, we continue using the row operations. First make all the leading entries one.

$$\begin{array}{ccc}
(-1/3)\rho_2 & x+y-z=2 \\
 & y-(2/3)z=5/3 \\
 & z=2
\end{array}$$

Finish by eliminating upwards with the leading entries.

Now we can read off the solution.

Finish by eliminating upwards with the leading entries.

Now we can read off the solution.

Using one entry to clear out the rest of a column is *pivoting* on that entry.

Example With this system

$$x-y -2w = 2$$

 $x + y + 3z + w = 1$
 $-y + z - w = 0$

we can rewrite in matrix notation and do Gauss's Method.

We can combine the operations making the leading entries to one.

Now eliminate upwards.

$$\stackrel{-(3/2)\rho_3+\rho_2}{\longrightarrow} \begin{pmatrix} 1 & -1 & 0 & -2 & 2 \\ 0 & 1 & 0 & 6/5 & -1/5 \\ 0 & 0 & 1 & 1/5 & -1/5 \end{pmatrix} \stackrel{\rho_2+\rho_1}{\longrightarrow} \begin{pmatrix} 1 & 0 & 0 & -4/5 & 9/5 \\ 0 & 1 & 0 & 6/5 & -1/5 \\ 0 & 0 & 1 & 1/5 & -1/5 \end{pmatrix}$$

The final augmented matrix

$$\begin{pmatrix}
1 & 0 & 0 & -4/5 & 9/5 \\
0 & 1 & 0 & 6/5 & -1/5 \\
0 & 0 & 1 & 1/5 & -1/5
\end{pmatrix}$$

leads to this description of the solution set.

$$\left\{ \begin{pmatrix} \frac{7/3}{-1/5} \\ -1/5 \\ 0 \end{pmatrix} + \begin{pmatrix} \frac{4/3}{-6/5} \\ -1/5 \\ 1 \end{pmatrix} w \mid w \in \mathbb{R} \right\}$$

Gauss-Jordan reduction

This extension of Gauss's Method is Gauss-Jordan reduction.

1.3 Definition A matrix or linear system is in reduced echelon form if, in addition to being in echelon form, each leading entry is a one and is the only nonzero entry in its column.

Gauss-Jordan reduction

This extension of Gauss's Method is Gauss-Jordan reduction.

1.3 Definition A matrix or linear system is in reduced echelon form if, in addition to being in echelon form, each leading entry is a one and is the only nonzero entry in its column.

The cost of using Gauss-Jordan reduction to solve a system is the additional arithmetic. The benefit is that we can just read off the solution set from the reduced echelon form.

Reduces to is an equivalence

1.5 *Lemma* Elementary row operations are reversible.

Reduces to is an equivalence

1.5 Lemma Elementary row operations are reversible.

Proof For any matrix A, the effect of swapping rows is reversed by swapping them back, multiplying a row by a nonzero k is undone by multiplying by 1/k, and adding a multiple of row i to row j (with $i \neq j$) is undone by subtracting the same multiple of row i from row j.

$$A \stackrel{\rho_i \leftrightarrow \rho_j}{\longrightarrow} \stackrel{\rho_j \leftrightarrow \rho_i}{\longrightarrow} A \qquad A \stackrel{k\rho_i}{\longrightarrow} \stackrel{(1/k)\rho_i}{\longrightarrow} A \qquad A \stackrel{k\rho_i + \rho_j}{\longrightarrow} \stackrel{-k\rho_i + \rho_j}{\longrightarrow} A$$

(The third clause is false if i = j.)

QED

Reduces to is an equivalence

1.5 *Lemma* Elementary row operations are reversible.

Proof For any matrix A, the effect of swapping rows is reversed by swapping them back, multiplying a row by a nonzero k is undone by multiplying by 1/k, and adding a multiple of row i to row j (with $i \neq j$) is undone by subtracting the same multiple of row i from row j.

$$A \stackrel{\rho_{\mathfrak{i}} \leftrightarrow \rho_{\mathfrak{j}}}{\longrightarrow} \stackrel{\rho_{\mathfrak{j}} \leftrightarrow \rho_{\mathfrak{i}}}{\longrightarrow} A \qquad A \stackrel{k\rho_{\mathfrak{i}}}{\longrightarrow} \stackrel{(1/k)\rho_{\mathfrak{i}}}{\longrightarrow} A \qquad A \stackrel{k\rho_{\mathfrak{i}} + \rho_{\mathfrak{j}}}{\longrightarrow} \stackrel{-k\rho_{\mathfrak{i}} + \rho_{\mathfrak{j}}}{\longrightarrow} A$$

(The third clause is false if i = j.)

QED

We say that matrices that reduce to each other are 'equivalent with respect to the relationship of row reducibility'. The next result justifies this using the definition of an equivalence. 1.6 Lemma Between matrices, 'reduces to' is an equivalence relation.

1.6 Lemma Between matrices, 'reduces to' is an equivalence relation.

Proof We must check the conditions (i) reflexivity, that any matrix reduces to itself, (ii) symmetry, that if A reduces to B then B reduces to A, and (iii) transitivity, that if A reduces to B and B reduces to C then A reduces to C.

1.6 Lemma Between matrices, 'reduces to' is an equivalence relation.

Proof We must check the conditions (i) reflexivity, that any matrix

reduces to itself, (ii) symmetry, that if A reduces to B then B reduces to A, and (iii) transitivity, that if A reduces to B and B reduces to C then A reduces to C.

Reflexivity is easy; any matrix reduces to itself in zero row operations.

The relationship is symmetric by the prior lemma—if A reduces to B by some row operations then also B reduces to A by reversing those operations.

1.6 Lemma Between matrices, 'reduces to' is an equivalence relation.

Proof We must check the conditions (i) reflexivity, that any matrix reduces to itself, (ii) symmetry, that if A reduces to B then B reduces to A, and (iii) transitivity, that if A reduces to B and B reduces to C then A reduces to C.

Reflexivity is easy; any matrix reduces to itself in zero row operations.

The relationship is symmetric by the prior lemma—if A reduces to B by some row operations then also B reduces to A by reversing those operations.

For transitivity, suppose that A reduces to B and that B reduces to C. Following the reduction steps from $A \to \cdots \to B$ with those from $B \to \cdots \to C$ gives a reduction from A to C. QED

1.7 Definition Two matrices that are inter-reducible by elementary row operations are row equivalent.

1.7 Definition Two matrices that are inter-reducible by elementary row operations are row equivalent.

The diagram below shows the collection of all matrices as a box. Inside that box, each matrix lies in some class. Matrices are in the same class if and only if they are interreducible. The classes are disjoint—no matrix is in two distinct classes. We have partitioned the collection of matrices into *row equivalence classes*.



Linear Combination Lemma

How Gauss's method acts

Example Consider this reduction.

$$\begin{pmatrix} 1 & 3 & 5 \\ 2 & 4 & 8 \end{pmatrix} \xrightarrow{-2\rho_1 + \rho_2} \begin{pmatrix} 1 & 3 & 5 \\ 0 & -2 & -2 \end{pmatrix} \xrightarrow{-(1/2)\rho_2} \begin{pmatrix} 1 & 3 & 5 \\ 0 & 1 & 1 \end{pmatrix}$$
$$\xrightarrow{-3\rho_2 + \rho_1} \begin{pmatrix} 1 & 0 & 2 \\ 0 & 1 & 1 \end{pmatrix}$$

Denote the matrices as $A \to D \to G \to B$ The steps take us through these combinations.

$$\left(\begin{array}{c} \alpha_1 \\ \alpha_2 \end{array} \right) \stackrel{-2\rho_1+\rho_2}{\longrightarrow} \left(\begin{array}{c} \delta_1 = \alpha_1 \\ \delta_2 = -2\alpha_1 + \alpha_2 \end{array} \right)$$

$$\stackrel{-(1/2)\rho_2}{\longrightarrow} \left(\begin{array}{c} \gamma_1 = \alpha_1 \\ \gamma_2 = -1\alpha_1 + (1/2)\alpha_2 \end{array} \right)$$

$$\stackrel{-3\rho_2+\rho_1}{\longrightarrow} \left(\begin{array}{c} \beta_1 = 4\alpha_1 - (3/2)\alpha_2 \\ \beta_2 = -1\alpha_1 + (1/2)\alpha_2 \end{array} \right)$$

How Gauss's method acts

Example Consider this reduction.

$$\begin{pmatrix} 1 & 3 & 5 \\ 2 & 4 & 8 \end{pmatrix} \xrightarrow{-2\rho_1 + \rho_2} \begin{pmatrix} 1 & 3 & 5 \\ 0 & -2 & -2 \end{pmatrix} \xrightarrow{-(1/2)\rho_2} \begin{pmatrix} 1 & 3 & 5 \\ 0 & 1 & 1 \end{pmatrix}$$
$$\xrightarrow{-3\rho_2 + \rho_1} \begin{pmatrix} 1 & 0 & 2 \\ 0 & 1 & 1 \end{pmatrix}$$

Denote the matrices as $A \to D \to G \to B$ The steps take us through these combinations.

$$\begin{pmatrix} \alpha_1 \\ \alpha_2 \end{pmatrix} \xrightarrow{-2\rho_1 + \rho_2} \begin{pmatrix} \delta_1 = \alpha_1 \\ \delta_2 = -2\alpha_1 + \alpha_2 \end{pmatrix}$$

$$\xrightarrow{-(1/2)\rho_2} \begin{pmatrix} \gamma_1 = \alpha_1 \\ \gamma_2 = -1\alpha_1 + (1/2)\alpha_2 \end{pmatrix}$$

$$\xrightarrow{-3\rho_2 + \rho_1} \begin{pmatrix} \beta_1 = 4\alpha_1 - (3/2)\alpha_2 \\ \beta_2 = -1\alpha_1 + (1/2)\alpha_2 \end{pmatrix}$$

Gauss's method systemmatically develops linear combinations of rows.

Linear Combination Lemma

2.3 *Lemma* A linear combination of linear combinations is a linear combination.

Linear Combination Lemma

.3 Lemma A linear combination of linear combinations is a linear combination.

Proof Given the set $c_{1,1}x_1 + \cdots + c_{1,n}x_n$ through $c_{m,1}x_1 + \cdots + c_{m,n}x_n$ of linear combinations of the x's, consider a combination of those

$$d_1(c_{1,1}x_1 + \cdots + c_{1,n}x_n) + \cdots + d_m(c_{m,1}x_1 + \cdots + c_{m,n}x_n)$$

where the d's are scalars along with the c's. Distributing those d's and regrouping gives

$$= (d_1c_{1,1} + \cdots + d_mc_{m,1})x_1 + \cdots + (d_1c_{1,n} + \cdots + d_mc_{m,n})x_n$$

QED

which is also a linear combination of the x's.

2.4 *Corollary* Where one matrix reduces to another, each row of the second is a linear combination of the rows of the first.

2.4 Corollary Where one matrix reduces to another, each row of the second is a linear combination of the rows of the first.

Proof We proceed by induction on the minimum number of row operations that take a first matrix A to a second one B. In the base step, that zero reduction operations suffice, the two matrices are equal and each row of B is trivially a combination of A's rows: $\vec{\beta}_i = 0 \cdot \vec{\alpha}_1 + \dots + 1 \cdot \vec{\alpha}_i + \dots + 0 \cdot \vec{\alpha}_m.$

$$\beta_i = 0 \cdot \vec{\alpha}_1 + \dots + 1 \cdot \vec{\alpha}_i + \dots + 0 \cdot \vec{\alpha}_m.$$

2.4 Corollary Where one matrix reduces to another, each row of the second is a linear combination of the rows of the first.

Proof We proceed by induction on the minimum number of row operations that take a first matrix A to a second one B. In the base step, that zero reduction operations suffice, the two matrices are equal and each row of B is trivially a combination of A's rows: $\vec{\beta}_i = 0 \cdot \vec{\alpha}_1 + \dots + 1 \cdot \vec{\alpha}_i + \dots + 0 \cdot \vec{\alpha}_m$.

For the inductive step, assume the inductive hypothesis: with $t\geqslant 0,$ any matrix that can be derived from A in t or fewer operations then has rows that are linear combinations of A's rows. Suppose that reducing from A to B requires t+1 operations. There must be a next-to-last matrix G so that $A\longrightarrow \cdots \longrightarrow G\longrightarrow B.$ The inductive hypothesis applies to this G because it is only t operations away from A. That is, each row of G is a linear combination of the rows of A.

If the operation taking G to B is a row swap then the rows of B are just the rows of G reordered, and thus each row of B is a linear combination of the rows of G. If the operation taking G to B is multiplication of some row i by a scalar c then the rows of B are a linear combination of the rows of G; in particular, $\vec{\beta}_i = c \vec{\gamma}_i$. And if the operation is adding a multiple of one row to another then clearly the rows of B are linear combinations of the rows of G. In all three cases the Linear Combination Lemma applies to show that each row of B is a linear combination of the rows of A.

With both a base step and an inductive step, the proposition follows by the principle of mathematical induction. QED

2.5 *Lemma* In an echelon form matrix, no nonzero row is a linear combination of the other nonzero rows.

2.5 *Lemma* In an echelon form matrix, no nonzero row is a linear combination of the other nonzero rows.

Proof Let R be in echelon form and consider the non- $\vec{0}$ rows. First observe that if we have a row written as a combination of the others $\vec{\rho}_i = c_1 \vec{\rho}_1 + \dots + c_{i-1} \vec{\rho}_{i-1} + c_{i+1} \vec{\rho}_{i+1} + \dots + c_m \vec{\rho}_m$ then we can rewrite that equation as

$$\vec{0} = c_1 \vec{\rho}_1 + \dots + c_{i-1} \vec{\rho}_{i-1} + c_i \vec{\rho}_i + c_{i+1} \vec{\rho}_{i+1} + \dots + c_m \vec{\rho}_m \quad (*)$$

where not all the coefficients are zero; specifically, $c_i=-1$. The converse holds also: given equation (*) where some $c_i\neq 0$ then we could express $\vec{\rho}_i$ as a combination of the other rows by moving $c_i\vec{\rho}_i$ to the left side and dividing by c_i . Therefore we will have proved the theorem if we show that in (*) all of the coefficients are 0. For that we use induction on the row index i.

The base case is the first row i=1 (if there is no such nonzero row, so R is the zero matrix, then the lemma holds vacuously). Recall our notation that ℓ_i is the column number of the leading entry in row i. Equation (*) applied to the entries of the rows from column ℓ_1 gives this.

$$0 = c_1 r_{1,\ell_1} + c_2 r_{2,\ell_1} + \dots + c_m r_{m,\ell_1}$$

The matrix is in echelon form so every row after the first has a zero entry in that column $r_{2,\ell_1} = \cdots = r_{m,\ell_1} = 0$. Thus $c_1 = 0$ because $r_{1,\ell_1} \neq 0$, as it leads the row.

The base case is the first row i=1 (if there is no such nonzero row, so R is the zero matrix, then the lemma holds vacuously). Recall our notation that ℓ_i is the column number of the leading entry in row i. Equation (*) applied to the entries of the rows from column ℓ_1 gives this.

$$0 = c_1 r_{1,\ell_1} + c_2 r_{2,\ell_1} + \dots + c_m r_{m,\ell_1}$$

The matrix is in echelon form so every row after the first has a zero entry in that column $r_{2,\ell_1}=\cdots=r_{m,\ell_1}=0$. Thus $c_1=0$ because $r_{1,\ell_1}\neq 0$, as it leads the row.

The inductive step is to prove this implication: if for each row index $k \in \{1, ..., i\}$ the coefficient c_k is 0 then c_{i+1} is also 0. Consider the entries from column ℓ_{i+1} in equation (*).

$$0 = c_1 r_{1,\ell_{i+1}} + \dots + c_{i+1} r_{i+1,\ell_{i+1}} + \dots + c_m r_{m,\ell_{i+1}}$$

By the inductive hypothesis the coefficients $c_1, \ldots c_i$ are all 0 so the equation reduces to $0 = c_{i+1}r_{i+1,\ell_{i+1}} + \cdots + c_m r_{m,\ell_{i+1}}$. As in the base case, because the matrix is in echelon form $r_{i+2,\ell_{i+1}} = \cdots = r_{m,\ell_{i+1}} = 0$ and $r_{i+1,\ell_{i+1}} \neq 0$. Thus $c_{i+1} = 0$.

Example In this non-echelon form matrix the third row is the sum of the first and second.

$$\begin{pmatrix} 1 & -1 & 3 \\ 2 & 0 & 4 \\ 3 & -1 & 7 \end{pmatrix}$$

Example In this non-echelon form matrix the third row is the sum of the first and second.

$$\begin{pmatrix} 1 & -1 & 3 \\ 2 & 0 & 4 \\ 3 & -1 & 7 \end{pmatrix}$$

But in the matrix that we get from Gauss's Method

the only linear relationship among the nonzero rows

$$\vec{0} = c_1(1 \quad -1 \quad 3) + c_2(0 \quad 2 \quad -2)$$

is the trivial relationship, since the equation of first entries $0 = c_1 \cdot 1$ gives that $c_1 = 0$ and then the equation of second entries $0 = 0 \cdot (-1) + c_2 \cdot 2$ gives $c_2 = 0$.

Proof Fix a number of rows m. We will proceed by induction on the number of columns n.

The base case is that the matrix has n=1 column. If this is the zero matrix then its unique echelon form is the zero matrix. If instead it has any nonzero entries then when the matrix is brought to reduced echelon form it must have at least one nonzero entry, so it has a 1 in the first row. Either way, its reduced echelon form is unique.

Proof Fix a number of rows m. We will proceed by induction on the number of columns n.

The base case is that the matrix has n=1 column. If this is the zero matrix then its unique echelon form is the zero matrix. If instead it has any nonzero entries then when the matrix is brought to reduced echelon form it must have at least one nonzero entry, so it has a 1 in the first row. Either way, its reduced echelon form is unique.

For the inductive step we assume that n>1 and that all m row matrices with fewer than n columns have a unique reduced echelon form. Consider an $m\times n$ matrix A and suppose that B and C are two reduced echelon form matrices derived from A. We will show that these two must be equal.

Proof Fix a number of rows m. We will proceed by induction on the number of columns n.

The base case is that the matrix has n=1 column. If this is the zero matrix then its unique echelon form is the zero matrix. If instead it has any nonzero entries then when the matrix is brought to reduced echelon form it must have at least one nonzero entry, so it has a 1 in the first row. Either way, its reduced echelon form is unique.

For the inductive step we assume that n>1 and that all m row matrices with fewer than n columns have a unique reduced echelon form. Consider an $m \times n$ matrix A and suppose that B and C are two reduced echelon form matrices derived from A. We will show that these two must be equal.

Let \hat{A} be the matrix consisting of the first n-1 columns of A. Observe that any sequence of row operations that bring A to reduced echelon form will also bring \hat{A} to reduced echelon form. By the inductive hypothesis this reduced echelon form of \hat{A} is unique, so if B and C differ then the difference must occur in their n-th columns.

Consider a homogeneous system of equations for which A is the matrix of coefficients.

$$a_{1,1}x_{1} + a_{1,2}x_{2} + \dots + a_{1,n}x_{n} = 0$$

$$a_{2,1}x_{1} + a_{2,2}x_{2} + \dots + a_{2,n}x_{n} = 0$$

$$\vdots$$

$$a_{m,1}x_{1} + a_{m,2}x_{2} + \dots + a_{m,n}x_{n} = 0$$
(*)

By Theorem One.I.1.5 the set of solutions to that system is the same as the set of solutions to B's system

$$b_{1,1}x_1 + b_{1,2}x_2 + \dots + b_{1,n}x_n = 0$$

$$b_{2,1}x_1 + b_{2,2}x_2 + \dots + b_{2,n}x_n = 0$$

$$\vdots$$

$$b_{m,1}x_1 + b_{m,2}x_2 + \dots + b_{m,n}x_n = 0$$
(**)

and to C's.

$$c_{1,1}x_{1} + c_{1,2}x_{2} + \dots + c_{1,n}x_{n} = 0$$

$$c_{2,1}x_{1} + c_{2,2}x_{2} + \dots + c_{2,n}x_{n} = 0$$

$$\vdots$$

$$c_{m,1}x_{1} + c_{m,2}x_{2} + \dots + c_{m,n}x_{n} = 0$$
(***)

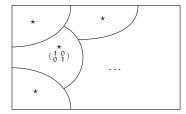
With B and C different only in column n, suppose that they differ in row i. Subtract row i of (***) from row i of (**) to get the equation $(b_{i,n}-c_{i,n})\cdot x_n=0$. We've assumed that $b_{i,n}\neq c_{i,n}$ so the system solution includes that $x_n=0$. Thus in (**) and (***) the n-th column contains a leading entry, or else the variable x_n would be free. That's a contradiction because with B and C equal on the first n-1 columns, the leading entries in the n-th column would have to be in the same row, and with both matrices in reduced echelon form, both leading entries would have to be 1, and would have to be the

QED

only nonzero entries in that column. Thus B = C.

With B and C different only in column n, suppose that they differ in row i. Subtract row i of (***) from row i of (**) to get the equation $(b_{i,n}-c_{i,n})\cdot x_n=0$. We've assumed that $b_{i,n}\neq c_{i,n}$ so the system solution includes that $x_n=0$. Thus in (**) and (***) the n-th column contains a leading entry, or else the variable x_n would be free. That's a contradiction because with B and C equal on the first n-1 columns, the leading entries in the n-th column would have to be in the same row, and with both matrices in reduced echelon form, both leading entries would have to be 1, and would have to be the only nonzero entries in that column. Thus B=C. QED

So the reduced echelon form is a canonical form for row equivalence: the reduced echelon form matrices are representatives of the classes.



Example To decide if these two are row equivalent

$$\begin{pmatrix} 3 & 2 & 0 \\ 1 & -1 & 2 \\ 4 & 1 & 2 \end{pmatrix} \qquad \begin{pmatrix} 3 & 1 & -2 \\ 6 & 2 & -4 \\ 1 & 0 & 2 \end{pmatrix}$$

use Gauss-Jordan elimination to bring each to reduced echelon form and see if they are equal.

Example To decide if these two are row equivalent

$$\begin{pmatrix} 3 & 2 & 0 \\ 1 & -1 & 2 \\ 4 & 1 & 2 \end{pmatrix} \qquad \begin{pmatrix} 3 & 1 & -2 \\ 6 & 2 & -4 \\ 1 & 0 & 2 \end{pmatrix}$$

use Gauss-Jordan elimination to bring each to reduced echelon form and see if they are equal. The results are

and

so the original matrices are not row equivalent.